The DoD

Software Tech News

WWW.SOFTWARETECHNEWS.COM

APRIL 2004 Vol.7. No.1

Grid Computing

Unclassified and Unlimited Distribution

http://iac.dtic.mil/dacs/

Tech View:

By Ellen Walker, DACS Analyst

Background

The idea for dedicating an issue of Software Tech News to the topic of Grid Computing was spurred by my involvement in a task for the Air Force Research Laboratory Information Directorate (AFRL/IF) Rome Research Site in 2003 to coordinate a Chief Scientist Guest Lecture Series that focused on "cyber infrastructure". My task was to identify potential topics and speakers and bring them to AFRL to present to the Lab's on-site personnel. Although I had heard the term "Grid Computing", I had a simplistic understanding of what it was. I had heard an analogy that likened it to the functions of the power companies providing power on demand, with users neither knowing nor caring where the power came from. I also had a few colleagues who were allowing their home PCs to be used for computing at night or during the day when they were at work and the computer would typically be idle. I knew that the researchers were attempting to process vast amounts of data by tapping into the resources of hundreds of thousands of PCs, each doing a small part. That was it – the extent of my knowledge.

Through research and attending the lecture series events I began to have my eyes opened to the magnitude and complexity of grid computing, and to the unique challenges of making independent autonomous resources actually work together as one virtual entity. I listened to speakers from IBM, and HP put forth their perspectives and discuss the benefits and challenges of Grid Computing. Lee Liming talked about the Globus Project, the key players, and the importance of standards. Charles Jung, from Internet2, talked about the NSF Middleware initiative and its contribution to Grid Computing. Patricia Kovatch talked about the TeraGrid and the operational aspects involved in making it a reality.

Grid computing is touted as the next evolution of the Internet that, together with other emerging technologies such as web services and virtual collaboration, will provide new ways to conduct business and research. Therefore, it seems appropriate to give our readers an opportunity to enhance their awareness of this emerging technology

Grid computing, by definition, cuts across organizational boundaries. Its development is supported by a multitude of working groups and forums where competitors often find themselves working together for the

continues on page 4

Grid Computing Defined

Grid computing is an emerging technology that transforms a computational infrastructure into an integrated, pervasive virtual environment for dynamic collaboration and shared resources no matter where they are located, and potentially anywhere in the world – providing users with unprecedented computing power, services and information.

Grids connect heterogeneous computing platforms and data sources so that they operate, and appear to the user, as a single computing system and data repository. Computational problems can be submitted to the grid for quick processing in the most cost effective manner – often faster and more efficient than monolithic supercomputers.

The Internet enables computers to communicate with each other. Grid enables computers to work together. Grid computing is the next evolution of the Internet dovetailing with the next generation of distributed computing, peer-to-peer computing, virtual collaboration technologies and Web services, resulting in new, more powerful ways to conduct business and research. Through the use of grids, businesses and researchers can complete computing-intensive tasks in a fraction of the time required with today's conventional resources.

***Source:** MCNC Grid Computing & Network Services* common goal and collaboration is the operative tool. This represents a significant paradigm shift for software development initiatives. Management of collaboration efforts is as important as solving the technical challenges.

In This Issue

The articles in this issue were selected to increase your awareness of the important technical issues involved in Grid computing and to convey a sense of its magnitude and complexity from both a management and technical perspective.

In the 1st featured article, titled "Grid Computing – a Services Perspective", Michael Osias comments on the origins and growth of Grid computing as well as its connection to the Global Information Grid (GIG) construct. He then focuses on Grid services, enabled by the development of web services technology, that allow the development of virtual systems that can be composed and re-configured dynamically at run time. Embedded throughout this very technical venue is the notion of sharing resources and services in a utility framework. Note also his reliance on standards.

The 2nd feature article, titled "The Role of Standards in the Grid", addresses the work of the Globus Alliance, the challenges they face, the importance of balancing short and long term needs, and their determination to have standards-based grids, with the accompanying rationale. It identifies and prioritizes standards work currently in progress, including the Globus toolkit, a product based on the Open Grid Services Architecture (OGSA), which is now used in most Grid implementations.

The 3rd feature article, titled "MCNC's North Carolina Statewide Grid Initiative", describes an actual Grid implementation. It represents a business perspective, focusing on both the actual and foreseeable benefits of using grid technology in addition to the specific challenges of that academic domain.

The last feature article, titled "TeraGrid Software Strategy: E Pluribus Unum", presents the operational reality of a grid computing environment from the perspective of those who must put it together and keep it functioning. I felt this article was extremely important to provide a realistic perspective to our technical readers. It also points to "new science" that is emerging given the capabilities of a grid infrastructure, and again, the use of standards based components.

I hope that the content of this newsletter will broaden your understanding of Grid computing, although it barely touches the surface of the technology. The DACS would be happy to provide follow-up material in a future issue, if you, our readers, so desire. We would like to hear from you. Please use our DACS website comment form at http/www.thedacs.com/forms/mailform.html to provide any comments about this issue, or requests for additional articles, or contact me.

Author Contact Information

Ellen Walker, DACS Analyst
Data and Analysis Center for Software
ITT Industries - Advanced Engineering and Sciences (AES)
775 Daedalian Dr.
Rome, NY 13441-490
Ph: (315)334.4936
Fax: (315)334.4964
E-mail: ellen.walker@itt.com

Grid Computing for Information Superiority

AFRL/IFTC
Advanced Computing Technology Branch
POC: Mr. Steve Drager and Dr. Lois Walsh/315-330-4029
iftc@rl.af.mil

Evolution in computing performance, capacity, and adaptability is necessary to continue advancing the ability of C4ISR systems to task, process, exploit and disseminate the voluminous information flowing from modern sensors and open sources. By increasing the computational capacity and sophistication of the architectural infrastructure, C4ISR modernization will be under girded to handle its mission realm of information acquisition, dissemination, exploitation and portrayal, decision-making and sense-making. While C4ISR systems are challenged by their complexity, these developments will confront the C4ISR complexity challenges and push "power to the edge."

Achieving, demonstrating, and fielding advanced computing architectures directly supports increased warfighter capabilities to support timely processing of information and maintenance of information superiority capability. Such capability is expected to result from the evolution in computing performance, capacity, interoperability and adaptability resulting from new technologies fused across the domains of High Performance Computing (HPC), embedded systems and information management. The development of grid computer architectures with greater capacity and sophistication will help to address dynamic mission objectives under constraints imposed by these systems in order to establish, maintain, and exploit information superiority.

The Joint Battlespace Infosphere (JBI) is being developed to provide the "right information" to the "right person" at the "right time." Scalability of the JBI as well as interactive real-time grid computing using High Performance Data Centers, which historically process batch submissions, are active areas of research.

The advanced computing group continues to explore ways to bring extramural and intramural researchers together for information sharing and development of possible collaborations. Individual visitors are welcome to discuss relevant research and possible collaborations.

Grid Computing — A Services Perspective

By Michael Osias, IBM

Grid Computing and Super Computing

Grid computing is distributed computing over the network enabled by open standards. It is the next evolutionary step in distributed computing resulting from the availability and convergence of capabilities offered by other distributed technology such as Common Object Request Broker Architecture (CORBA), distributed computing environment (DCE), Java 2 Enterprise Edition (J2EE), Web services, and the build out of network capacity during the late 90s. The increasing computing demands and requirements that governments, business, and science and research are placing on information technology are driving the convergence of these distributed technologies to address these demands.

Initially, Grid computing was formed in the scientific and research communities to enable large scale computations whose capacity requirements exceeded any single organization's resources. Collaborating researchers needed a way to aggregate and share super computing resources and make them available to do their work. In essence, by linking these resources together using Grid resource sharing protocols, a larger, virtual super computer was formed composed of the shared resources. Often, these shared resources were comprised of heterogeneous compute and storage resources and tended to be geographically distributed. Increasingly, industry began to face the same pressures and needs of having to handle increased workload and data storage requirements, often within or between distributed enterprises, without additional capacity.

As these types of large scale shared resource complexes became more commonplace, efforts to standardize the infrastructure and services began. Early leaders of the standardization efforts were from research, science, and academia. Industry and governments, searching for solutions to similar problems, recognized the emerging Grid technology as a viable computing technology to solve problems today and provide a clear evolutionary path to the future. As adoption grew, so did the participation in the development of the standards, which today include participants from industry, government, and academia. As a result of this expanded adoption and participation, there is a rich set of Grid computing standards and technology.

As these types of large scale shared resource complexes became more commonplace, efforts to standardize the infrastructure and services began.

Industry leaders are heavily involved in developing and refining standards. IBM in particular has played a significant role in the development of the standards and contributed to the development of the open source standards based software, the Globus Toolkit. IBM has incorporated Grid technology across all lines of business including services, hardware, and software, and is delivering over 100 leading edge production Grid projects around the world.

Grids of a Feather: The Global Information Grid and Grid Computing

Grid computing is a parallel and distributed architecture that enables the sharing, selection, and aggregation of geographically distributed resources. Specific kinds of resources applicable to the Department of Defense (DoD) environment include computers (PCs, workstations, clusters, supercomputers, laptops, notebooks, mobile devices, PDAs, etc), software (e.g., Service Providers vending content and expensive special purpose applications on demand), catalogued data and databases (e.g. transparent access to unit readiness, weapons characteristics databases etc.), special devices and sensors (e.g., reconnaissance satellites, UAVs, etc.). Additionally, Grid technology addresses resource availability, capability, costs, and user Quality of Service (QoS) requirements. Grid infrastructure services, such as schedulers, resource managers, information services, and data services, are all essential to the Grid complex to provide inherent resiliency, resource optimization, and non-trivial quality of service while managing resources in the battlespace or DoD environment.

Traditionally, as a reflection of the origin of Grid technology, the usage model is oriented toward efficient use of all available processing horsepower. Grid standards are now leveraging innovations in Web services technology with Open Grid Services Architecture (OGSA), moving Grid technology toward a service-oriented computing

continues on page 6

Grid Computing – A Services Perspective

Continued from page 5.

model. This enables market-based utility computing and a new Net-Centric paradigm.

Transforming defense intranets into powerful, self-managing, virtual computers is a technical objective that Grid technology can enable. This large scale virtual computing environment, supported and enabled by a common Grid computing substrate, enables a processing environment for full Net-Centricity. It fits with the DoD Global Information Grid (GIG) vision, with DoD's current state consisting of a vast collection of heterogeneous systems, just beginning to share communications and computing resources on wide area networks. Soon the GIG will be moving to implement the GIG Enterprise Services (GES). These enterprise services provide ubiquitous service-oriented access to DoD data. A Grid complex supporting these data services means that they perform according to quality of service policies, are standards based, and universally accessible to users and other applications. Grid services technology allows rapid, efficient coupling of legacy transactions systems and will enable new warrior or business applications. In addition, Grid-enabled enterprise services will begin to address globally distributed computing issues.

Grid Proliferation and the Global Grid

As the momentum of Grid adoption continues to grow, we will see more domains and more Grids being built out within, and at times, between these domains. The proliferation of Grid technology is where the Grid standards are so important. To be able to join together Grids, whether the intent is to tap into another organization's Grid for collaboration and form an inter-Grid, to interface with an infrastructure provider to augment an existing infrastructure and dynamically increase capacity on demand, or if statewide Grids want to interconnect with other states, universities, or federal Grids, the open Grid standards enable Grid interconnectivity. As the technology matures and Grids continue to be built out, inter-Grids will become more commonplace and provide the technical and architectural substrate for the next generation of computing.

> *Transforming defense intranets into powerful, self-managing, virtual computers is a technical objective that Grid technology can enable.*

In addition to allowing organizations to share resources within or across organizational boundaries in an inter-Grid scenario, individuals or departments may want to make resources available to a larger Grid. Standards again are the enabler, and using standards-based Grid middleware infrastructure services, resources can easily be added to existing Grids and made available for use. There are many choices for this infrastructure software, from open source software such as the J2EE-based Globus Toolkit to commercial offerings such as the IBM Grid Toolbox and DataSynapse.

Getting Started with Grid Services

Web services technology provides a standard approach to separating interface from implementation details in an Internet-based distributed computing environment. By decoupling interface from implementation, classical Grid technology can be used to develop systems that can be composed and reconfigured dynamically at run time. Building on Web services technologies, the Open Grid Services Infrastructure (OGSI) specification defines Web service interfaces and behaviors for creating, managing, and exchanging information among stateful Web service instances. OGSI is comprised of an extension of the base Web services standards, which currently do not address the notion of stateful Web services instances, and the definition of a set of interfaces that describe and standardize requestor-service interactions. These interactions provide the basis for additional semantic interoperability not provided in the base Web service standards. In essence, these interfaces define the basis of a Web services programming model.

With the exception of the simplest data retrieval or computation services, a service needs to keep some kind of state to provide any value above and beyond a basic Web-based extensible markup language (XML) document processor or a remote procedure call (RPC) interface. Many services have a distinct state associated with every single client accessing them. A common example of such a service is a shopping application with the shopping cart unique for each client. An e-commerce business can be created simply by integrating and coordinating sets of services which include on demand supply and procurement, shipping and routing, and accounting and billing, as well as Web application infrastructure services such as a shopping cart, an online address book service, and payment processing. By using available services, like a generic shopping cart service, the company saves money by assembling these component services into higher level services and applications. There is no need for them to develop and manage code and infrastructure for basic features such as a shopping cart. The service provider that provides the shopping cart must be able to accommodate many different shopping carts from multiple customers. In addition to the obvious service data elements needed for state such as the shopping cart contents and session information, the service provider must also track activity and usage data for billing the appropriate organization (not necessarily the person using the shopping cart, but the e-commerce customer who has integrated the service into their online business).

Another example is that of a database query service. It is analogous to using a database command client, such as a DB2 command line interface, where each client has its own connection and consequently connection session state to the database server. This allows concurrent access to the database server with different clients performing a variety of database operations simultaneously. To extend the notion of a database connection to a services oriented architecture, a similar session and state management facility is required. Otherwise, using the classical Web services definition and forcing an analogy of a Web service as a singleton, the service is able to accommodate only a single data service client at a time, typically in the call pattern of a blocking call. Using the OGSI component programming model, each client can have its own instance of a service by either inspecting a Grid data service registry or by invoking the createService() port reference to instantiate a new service instance. With this model, multiple clients can interact with the database service with their own individual session state preserved. An implementation strategy for this kind of service could involve exposing stateful j2ee entity and session beans as ogsi-compliant Grid services to provide data access.

A reliable file transfer service, like the one provided with the open source Globus Toolkit, also requires the ability to maintain instances and state if it is to service more than a single client at any given time. The persistent factory service again is used to create a unique instance of the service for a client. For those familiar with software patterns, the factory pattern mentioned herein is very similar in concept. It is an OGSI base service whose purpose is to create instances of services, much like a static Java factory object would create well formed instances of specific objects using getInstance(). Once a service instance is created, the client can then interact with the service to provide such information as source and destination, and even register for periodic updates from the service instance. The service state is maintained by service data elements which are stored as part of the service instance. Each instance has its own set of service data which has different values. An example of the service data for the reliable file transfer service is the percentage of the file that has been transferred. A client may subscribe to this service data element using the OGSI notification framework to receive periodic updates about the amount of the file transferred. By having stateful service instances, potentially transient but also with a lifetime beyond that of the initial SOAP envelope processing, a Grid service can provide services to many different clients, and possibly offer different quality-of-service levels and corresponding implementations based on the priority of the user. As an example a Grid service may manage bandwidth allocation between two network endpoints based on the user level and priority, allocating a higher percentage of the link to the higher privileged user.

The OGSI programming model is designed to integrate key Grid technologies to create a distributed system framework to provide for the controlled management of the distributed and often long-lived state that is commonly required in sophisticated distributed applications. Fundamental Grid technologies include information services, resource management, data services, and security. The creation of Grid service instances using a persistent factory service enables the leveraging of the resources in terms of intelligent load balancing and service routing. A createService operation on a factory service may involve a number of steps before the service implementation is instantiated and a Grid Service Handle (GSH) returned to the client. For example, the factory service may intercept the instantiation of the service implementation to determine if a suitable instance already exists. It may also inspect the Grid infrastructure by using OGSA-managed resource services to determine the current load on a node, or consult an index information service to determine what suitable resources are available in the Grid complex. Ultimately, the factory may instantiate a new hosting environment on another node which is used to instantiate the service. A GSH for the service instance on the new node is registered in the index service and returned to the client. The factory service may have its own mechanisms for these processing cases, or may hand off to underlying resource management Grid infrastructure services such as brokers and schedulers. The factory service then becomes more of a coordinator of workflow across various Grid infrastructure services.

There are many resources such as IBM RedBooks and the IBM DeveloperWorks website Grid Computing zone. Also available is the IBM Grid Toolbox for Multiplatforms v3.0. This enhanced derivative of Globus allows developers to easily install, configure, develop, and manage Grid services. Installation, systems management, and application development enhancements let developers focus on Grid applications

continues on page 8

Global Information Grid – Questions and Answers

By Ricardo Bueno, ITT Industries, Advanced Engineering Sciences Division

What is the relationship of Grid Computing and the Global Information Grid - a few statements about how the government and military seek to implement grid computing?

The Global Information Grid is defined as a "Globally interconnected, end-to-end set of information capabilities, associated processes, and personnel for collecting, processing, storing, disseminating, and managing information on demand to warfighters, policy makers, and support personnel. The GIG includes all owned and leased communications and computing systems and services, software (including applications), data, security services, and other associated services necessary to achieve Information Superiority. It also includes National Security Systems (NSS) as defined in section 5142 of the Clinger-Cohen Act of 1996. The GIG supports all DoD, National Security, and related Intelligence Community (IC) missions and functions (strategic, operational, tactical, and business) in war and in peace. The GIG provides capabilities from all operating locations (bases, posts, camps, stations, facilities, mobile platforms, and deployed sites). The GIG provides interfaces to coalition, allied, and non-DoD users and systems.[1]

Grid computing is still considered an infant technology and its applications vary widely in use. Grid computing is predominantly used in the scientific and academic communities but of late much attention has been paid to what it might offer in information dominance and superiority.

There are a variety of uses for grid computing, each with a specific and significant capability, together providing the promise of unparalleled access to computing and information resources. The three types of Grid computing are; Computations Grids, Data Grids, and Foraging Grids. The Department of Defense is currently defining the GIG Enterprise Services (ES) under the Data Grid type with its appropriate architecture. The Defense Information Systems Agency (DISA), in collaboration with the joint services, is responsible for providing enterprise level services for the warfighter, a large community of users. These services, know as the Net-centric Enterprise Services or NCES, provide capabilities in the GIG ES such as distributed storage, information assurance/security, collaboration, messaging, mediation, service management, and interfaces for

continues on page 9

1 Global Information Grid Capstone Requirements Document DTD 30 Aug 2001

Grid Computing – A Services Perspective

Continued from page 5.

development and less on the Grid infrastructure itself.

About the Author

Mr. Michael J. Osias is a Technical Architect for IBM Grid Computing Americas, with a focus on Department of Defense and other Government agencies. He has e-Business technical architecture experience in large e-Business infrastructures including portals, B2B, e-Commerce, and legacy integration and transformation. Michael has a background in telecommunications, treasury and finance operations, and defense industry. Prior to being a Grid Technical Architect, Michael was an e-Business Architect covering several telcom and manufacturing accounts, while conducting independent Grid Computing research and participating in the Global Grid Forum. He received an IBM Invention Achievement Award for patent submissions and technical disclosures in such areas as web search technology and distributed transaction processing. Also, he is an IBM Certified WebSphere Systems Expert. He has an educational background in Mathematics and Computer Science with the State University of New York and other higher education institutions.

Author Contact Information
Michael J. Osias
E-mail: mosias@us.ibm.com
Cell: (201) 230-4905
URL: http://www.ibm.com/grid

Career Planning

Those looking to become key players in Grid technology and get started developing Grid services and Grid architectures should have a background in J2EE and Web services technologies, as well as distributed systems. It's a small step from there to Grid computing and Grid services.

node/client connectivity. Core management disciplines within the GIG ES include, among others, information engineering and management, applied metadata standards for information interoperability, and complex policy assignment for multinational/coalition forces. Over the next several years the GIG ES working groups will work hard to define and refine the architecture and functionality of the GIG ES.

What is the status/reality of the Global Information Grid?

The GIG ES is a real thing. Currently there have been defined 9 core services. These services are Storage, Messaging, Enterprise Service Management, Discovery, Mediation, Information Assurance, Application Hosting, User Assistant, and Collaboration. Each one of these core services is supported by a community of engineers and architects from all branches of the service. DISA is the executive agent for the GIG ES and is responsible for publishing the core services requirements as defined by the joint engineering and architecture communities. Furthermore, DISA is currently going through a detailed process for selecting Alternative of Analysis programs for GIG ES implementation. Some of the candidates for GIG ES are the Air Force's Global Combat Support System (GCSS-AF) and the Joint Battlespace Infosphere (JBI). Much work must still be done to ensure these programs fit the needs of the DoD and the requirements of the GIG ES.

Is it still someone's dream or is it operational?

While the GIG ES is not operational it is, nevertheless, far from being just a dream. The GIG ES program is a living, breathing, growing program that will have far-reaching and significant impact and benefit to the warfighter community. While not yet in use, the first set of prototype demonstrations is set for sometime in 2007.

About the Author

Ricardo Bueno serves on the following working groups for GIG ES supporting AF/XI and AF/CIO.

Metadata Working Group

Information Security Working Group

GIG ES Architecture Working Group

Data and Storage Working Group

He is also a core reviewer for the AF of the Initial Capabilities Document (ICD) which defines the core capabilities required for the GIG and the Capability Development Document (CDD) which provides details on acquiring and use cases for the capabilities defined in the ICD.

Contact Information

Ricardo A. Bueno, *Principal Systems Architect*
Joint Battlespace Infosphere Program
ITT Industries, Advanced Engineering and Sciences
Air Force Research Laboratory/IFSE
DSN 587
Phone:(315)330-2284
Cell: (315)335-1229
Fax: (315)330 -7009
E-mail:ricardo.bueno@rl.af.mil

The Role of Standards in the Grid

By Ian Foster and Lee Liming, Argonne National Laboratory

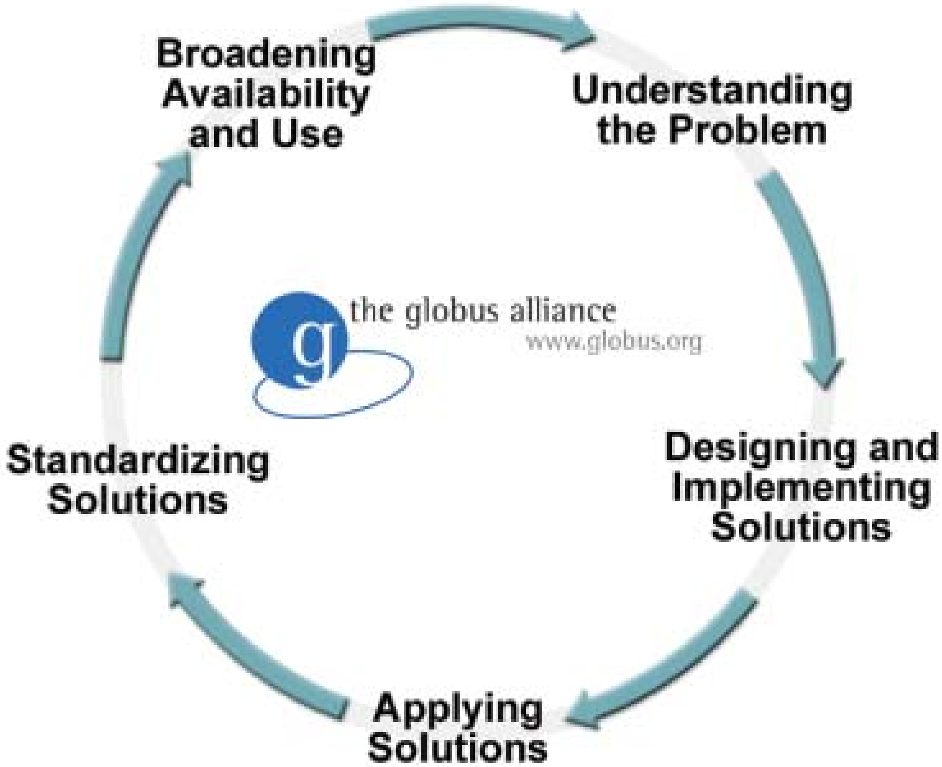


Figure 1. The Globus Alliance employs an iterative design methodology for the development and evaluation of proposed Grid technologies.

The Grid is a set of technologies that, when deployed across an enterprise's resources, ease the creation of applications and infrastructure that facilitate *collaborative* and *adaptive* systems.[1] In this article, we will explore the vital role that standards play in establishing the Grid. We also discuss the important role played by open software, in particular the software developed within the Globus Alliance, a collaboration of researchers, system architects, and software developers that has been pursuing this goal for several years.

Developers of Grid systems can build on a range of established and proposed standards that have emerged from work on the Internet, the World Wide Web, and Web-based applications—technologies that, like the Grid, seek to support resource sharing and collaboration. However, a number of key capabilities must be added before interoperable Grids can be constructed. Much work is currently underway to develop specifications, implement solutions, and propose standards (or additions to existing standards) for Grid computing. A particular focus of these efforts is system monitoring and management, an area that has not seen the broad adoption of standard solutions analogous to the Internet and the World Wide Web.[2] This area has instead seen an abundance of proprietary systems and specifications without any particular set of solutions becoming dominant.

The Work of the Globus Alliance

Our work in the Globus Alliance has historically consisted of the following activities, each of which represents a key challenge in the establishment of a useful Grid.

- Identifying and refining our understanding of the fundamental problems shared by potential users of the Grid (those who would develop or use Grid-based applications)
- Developing and refining solutions (including both specifications and implementations) to these fundamental problems
- Applying these solutions to real problems in order to test their applicability
- Proposing and championing these solutions as standards in broader communities
- Expanding availability (implementations and support) and adoption (uses in many settings) for these standard solutions

Our approach to realizing the Grid vision creates an ongoing tension between the short-term (but ongoing) need to have working software that supports user requirements and the longer-term need to formulate and promote new solutions that "push the envelope" of capabilities. In these early stages of the "Grid era"—when research is ongoing and constantly changing—this tension is quite strong: the solutions are evolving quickly enough that working implementations cannot keep up without making some sacrifices in continuity and backwards-compatibility. Our challenge is to find a good balance that serves both needs reasonably well.

Another challenge faced by the Globus Alliance (and the Grid

continues on page 11

community at large) is the need to formulate its solutions and standards in ways that appeal to the larger IT industry and the market it serves. Failure to meet this challenge would result in Grid computing languishing as a niche market, almost certainly without the critical mass required to sustain itself over time. Until standard Grid solutions are incorporated into mass market infrastructure technologies (as the Internet and the Web have been integrated into every major computer system now sold—not to mention cellular telephones, televisions, automobiles, and other commodity products), the Grid will remain beyond the reach of most of the users it is intended to help.

The Importance of Standards

The Grid vision requires protocols (and interfaces and policies) that are not only open and general-purpose but also *standard*. Indeed, we would argue that "Grids will be standards-based or not at all." It is standardization that allows potential collaborators to establish resource-sharing arrangements quickly and easily with *any* interested party. It is standardization that will allow organizations to establish resource-sharing contracts routinely for acquiring resources on demand, thus avoiding the need to build expensive data centers designed to handle peak loads that remain underutilized most of the time. Standard solutions can allow us to move away from today's plethora of balkanized, incompatible, non-interoperable distributed systems and toward a model where computing and data capabilities are available as standardized, interchangeable commodities.

The World Wide Web Consortium (W3C) has been highly successful at promoting a number of proposed standards in various stages of development that are commonly referred to as the "Web services" specifications.[3] The bases for these standards are the HTTP protocol, the XML encoding format, the SOAP remote procedure call mechanism, and the WSDL language. Web services standards have already been adopted by every major provider in the IT industry, and it is clear that they are heading toward a long life and a ubiquitous presence in current and future IT systems.

Until standard Grid solutions are incorporated into mass market infrastructure technologies . . . the Grid will remain beyond the reach of most of the users it is intended to help.

While Web services standards meet many needs in industry and are highly popular, there are currently no established Web services standards for the interfaces used to manage, monitor, and interact with resources and with services that maintain persistent state. Thus, it is not possible to define tools that monitor, manage, troubleshoot, etc., diverse resources in standard ways: critical issues if one is to build large-scale systems. It is precisely these issues with which Grid computing is concerned.

Our current strategy, therefore, is to ensure the establishment of Web services-based standards in these areas, and the Globus Alliance's efforts are currently focused on contributing significantly to this work. Our goal is a convergence of new and existing Web services standards with the solutions developed by the Grid community. A further goal is the standardization of the resulting technology so that it becomes at least as accessible as current Web services technology.

Accomplishments

Standards organizations with which we are working to formalize specifications for the Grid include the Global Grid Forum (GGF)[4] , the World Wide Web Consortium (W3C), the Organization for the Advancement of Structured Information Standards (OASIS)[5] , and the Internet Engineering Task Force (IETF).[6]

The Global Grid Forum has played a key role in developing and articulating the Open Grid Services Architecture (OGSA), which defines the Grid community's "guiding principles" for Web services/Grid convergence.[7] The GGF Open Grid Service Infrastructure (OGSI) Working Group formalized the OGSI v1.0 specification in 2003, which was our first attempt to define Web services standards for Grid computing.[8] The open source Globus Toolkit® 3.0 (released in mid-2003) contained the first OGSI implementation and was followed soon after by a number of independent implementations (see sidebar).[9]

We were disappointed (but not surprised) to learn subsequently that the Web services community was not satisfied with OGSI v1.0, due largely to a belief that OGSI v1.0 did not fit well enough with existing Web services architectural principles. In late 2003 and early 2004, a smaller group of Grid and Web services architects (leaders in their respective fields) met intensively to work out a compromise set of specifications, which was announced at the GlobusWORLD 2004 conference in January 2004. This set of specifications is known collectively as the Web Services Resource Framework, or WSRF.[10] It is our belief, expressed by leaders of both the Web services and Grid communities, that WSRF is the much-desired convergence of Web services and Grid technologies and that it will lead to a body of widely-available, highly-interoperable Grid technology. Of course, work to finalize the now-

continues on page 12

The Role of Standards in the Grid

Continued from page 11.

drafted WSRF specifications and to gain formal acceptance and endorsement by one or more standards bodies is yet to be done. The Globus Toolkit 4.0, expected in mid-2004, will include a WSRF implementation.

While WSRF is undoubtedly a major milestone in the Grid's history, it is nevertheless only one of many standards-related activities undertaken by the Grid community and by the Globus Alliance. The IETF recently approved our X.509 Proxy Certificate Profile as a proposed standard. This action standardizes the format of the Grid "proxy" certificates used to support the Grid's essential single sign-on and delegation capabilities. The GGF recently accepted our specification of the GridFTP v1.0 protocol for high-bandwidth data transfer over wide area networks.

We have also been working in the Grid and Web services communities toward draft specifications for Data Access and Integration (DAI) in the GGF, a WS-Agreement specification in the W3C for establishing service agreements between service providers and consumers (e.g., reserved network bandwidth or compute node scheduling priorities), and replica location capabilities for data management systems. We have actively engaged ongoing standards work in other areas, making Grid-focused contributions to the development of the WSDL 2.0 (Web Service Description Language) specification, the WSDM (distributed web service management) specification, and the SAML (Security Assertion Markup Language - an XML-based framework for exchanging security information) and XACML (Extensible Access Control Markup Language - an XML specification for expressing policies for information access over the Internet) languages for encoding policy specifications.

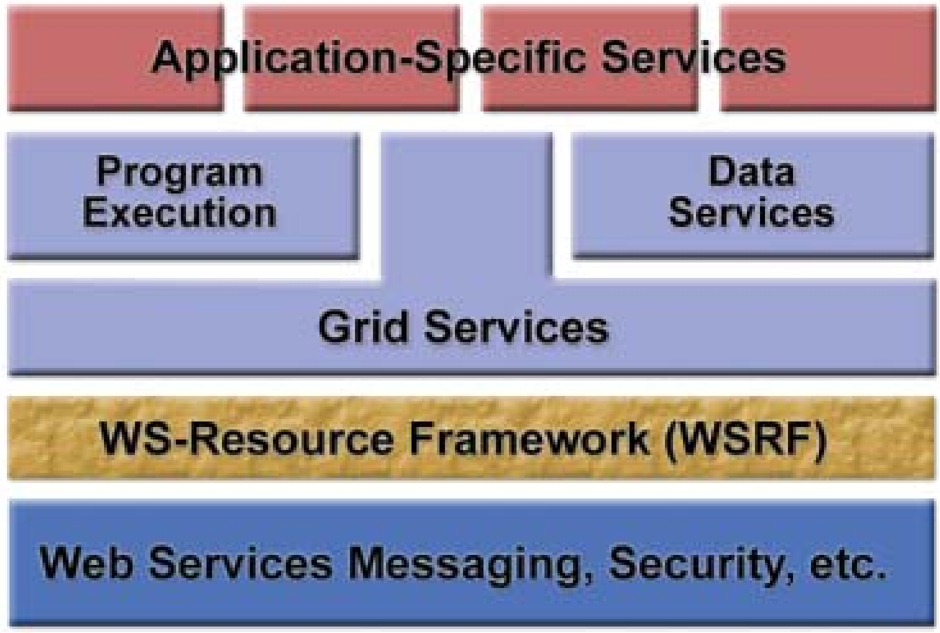


Figure 2. The Web Services Resource Framework provides a layer of commonality for all Grid Services based on the Web services model.

Conclusions and Invitations

In our view, the definition and widespread adoption of standard protocols and interfaces is currently the single most critical problem facing the Grid community. Fortunately, we are making good progress. On the standards side, we have the increasingly effective Global Grid Forum and representation in other IT standards organizations. We have broadly-endorsed efforts underway to define OGSA and WSRF, which enshrine Grid technology within the highly-successful Web services suite of products and solutions.

On the implementation side, seven years of experience and refinement have produced a widely used *de facto* standard, the open source Globus Toolkit. IBM, Microsoft, Platform, HP, Sun, Avaki, Entropia, United Devices, and other IT industry members have expressed strong support for OGSA and announced their own products and services based on OGSA standards. In time, we will be able to state that for an entity to be part of *the* Grid it must implement OGSA protocols and interfaces, just as to be part of the Internet an entity must speak IP (among other things). Both open source and commercial products will interoperate effectively in this heterogeneous, multi-vendor Grid world, thus providing the pervasive infrastructure needed to support compelling Grid applications.

Establishing the Grid and the standards that define it is still very much a "work in progress." Because we are taking a standards-based approach to the problem, it is open to input from many sources. The standards organizations (GGF, W3C, OASIS, IETF) all have open memberships, allowing organizations and individuals to become contributors to their work. The Globus Alliance also welcomes new partners in our broad mission to realize the Grid vision through the Globus Alliance Affiliates program (for academic and R&D members) and the Globus Alliance Commercial Affiliates program (for corporate participation). See www.globus.org for details.

continues on page 13

References

1 I. Foster. What is the Grid? A Three Point Checklist. *GRIDToday*, July 20, 2002.

2 I. Foster, C. Kesselman, S. Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International J. Supercomputer Applications*, 15(3), 2001.

3 World Wide Web Consortium (W3C), http://www.w3.org/. The W3C's activities with respect to Web services are detailed at http://www.w3.org/2002/ws/.

4 Global Grid Forum, http://www.grid-forum.org/.

5 Organization for the Advancement of Structured Information Standards (OASIS), http://www.oasis-open.org/.

6 Internet Engineering Task Force (IETF), http://www.ietf.org/.

7 I. Foster, C. Kesselman, J. Nick, S. Tuecke. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration. Open Grid Service Infrastructure WG, Global Grid Forum, June 22, 2002.

8 S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maguire, T. Sandholm, P. Vanderbilt, D. Snelling. Open Grid Services Infrastructure (OGSI) Version 1.0. Global Grid Forum Draft Recommendation, 6/27/2003.

9 The Globus Toolkit software is described in detail and available for download at http://www.globus.org/toolkit/. Globus Toolkit is a registered trademark held by the University of Chicago.

10 WSRF status and activities are described in detail at http://www.globus.org/wsrf/.

Standards-based Grid Products

Although Grid standards are still being formalized, a number of commercial and open source products that comply with early versions of these standards are available today. The first WSRF implementations are not expected until mid-2004, but many OGSI implementations already exist and conversion to WSRF is expected to follow.

Open Source Products

The Globus Toolkit is an open source software toolkit used for building standards-based Grids. It includes reference implementations of the emerging Grid standards, as well as tools for managing grid systems and developing Grid applications. Work on the Globus Toolkit is coordinated by the Globus Alliance, while many organizations have contributed to its components and features. See www.globus.org.

The NSF Middleware Initiative (NMI) software distribution includes a wide range of software, services, documents and recommendations for the effective use of information technology in research and education. Included are a number of popular open source Grid software tools. See www.nsf-middleware.org.

The Virtual Data Toolkit (VDT) includes Grid software designed for scientific and engineering projects that utilize very large (petabyte-scale) datasets. See www.griphyn.org/vdt.

The Open Grid Services Infrastructure (OGSI) extends Web services to meet several key Grid requirements. See www.globus.org/ogsa. OGSI implementations are available in a variety of languages, including Java (Globus Toolkit 3.0, Fujitsu Unicore), .NET (U. Virginia' s OGSI.NET and U. Edinburgh), Python (Lawrence Berkeley National Laboratory) and Perl (U. Manchester).

Commercial Products

IBM's WebSphere Application Server provides support for developing and deploying OGSA-based Grid services. IBM's Grid Toolbox provides implementations of standard Grid services and tools as well as additional tools and services for deploying Grid technologies within or across organizations. See www.ibm.com/grid.

HP offers a commercial version of the Globus Toolkit for HP-UX, Tru64 Unix, and Linux. The HP Utility Data Center (UDC) product line includes Data Center operation & construction services and HP's stated intent is to construct UDC products using OGSI-compliant services. HP also offers Enterprise Grid Consulting from HP Services. See www.hp.com/techservers/grid.

Platform offers a commercial version of the Globus Toolkit for Tru64, HP-UX, AIX, Linux, IRIX and Solaris, including integration with popular high-performance scheduling systems. Platform also offers Grid consulting and training from Platform Professional Services. See www.platform.com/grid.

About the Authors

Lee Liming is the Manager of the Distributed Systems Laboratory at Argonne National Laboratory, the "Chicago home" of the Globus Alliance. His work with Ian Foster, Carl Kesselman, and Steve Tuecke on Grid computing began in 1999, and he is currently responsible for work resulting in the Globus Toolkit and related Grid technologies. Through his work at Argonne National Laboratory, Mr. Liming is engaged in a broad range of Grid deployment and application projects including NEESgrid, the NSF Middeware Initiative/GRIDS Center, the NASA Information Power Grid, and the National Computational Science Alliance. He was previously a product manager and software engineer at ProQuest Information and Learning and an IT manager at the University of Michigan's School of Information.

Having joined the Mathematics and Computer Science Division of Argonne National Laboratory in 1989 as a postdoctoral researcher, **Ian Foster** has risen to the rank of senior scientist and associate division director. Since 1996, he has held a joint appointment in the department of computer science at the University of Chicago, where he now has the rank of professor. His research specifically addresses scalable authentication and authorization technologies (the Grid Security Infrastructure), information discovery and dissemination (the Meta Directory Service), resource reservation and management, high-performance data transfer, and data replication. He

continues on page 14

MCNC's North Carolina Statewide Grid Computing Initiative

Deploying Emerging Technologies to Spark Innovation and Drive Technology-Based Economic Development

By Phil Emer, Chuck Kesler and Lavanya Ramakrishnan, MCNC

MCNC Grid Computing & Networking Services, a member of the MCNC non-profit independent family of companies located in North Carolina's Research Triangle Park, and the University of North Carolina's 16-campus system are grid-enabling the existing statewide research and education network that interconnects universities throughout North Carolina.

This statewide grid is anticipated to be the first statewide research and education grid in the country. The initiative is viewed as the most ambitious upgrade to the state's computing infrastructure in history and a catalyst for economic development. It will support multiple scientific disciplines in addition to other grid information technology applications, such as administrative and library services.

The initiative emerged from MCNC's development of one of the country's first scientific computing grid networks in 2001, the North Carolina Bioinformatics Grid Test Bed, and has now been expanded to include:

- The MCNC Enterprise Grid
- The statewide grid network
- The development of a Grid Technology Evaluation Center

Details of these initiative subsets are provided in the following sections of this article.

Grid computing will benefit urban and rural areas of the state, spanning business, academia and government. It is especially important to smaller institutions that only need computing resources periodically and often cannot afford to invest in new technologies.

Most of the state's high-performance computing resources are at the large research universities surrounding Research Triangle Park – the University of North Carolina at Chapel Hill, Duke University and N.C. State University. Smaller universities, many in the more rural areas of the state, have historically lacked access to advanced computing resources. By enabling researchers and educators throughout the state to take advantage of computing resources that already exist at the large universities, and enabling all researchers in the state to pool resources and intellectual expertise, resources available to an individual researcher anywhere in the state will vastly increase. The statewide grid will be a catalyst for greater levels of innovation, the creation of more intellectual property and ultimately more businesses started with local entrepreneurial leadership throughout North Carolina.

The Foundation of the Statewide Grid

The North Carolina Research and Education Network (NCREN), established in 1985 through a collaboration of MCNC and the University of North Carolina system, is the backbone infrastructure for the statewide grid.

Operated by MCNC, it has evolved along with the Internet from a research project to a critical infrastructure for the research and education community. NCREN is a production-level, IP (Internet Protocol) network providing advanced communications and Internet services to more than 180 locations, including universities and other government and non-profit institutions throughout North Carolina. It serves about half a million students, faculty and staff from the University of North Carolina's 16-campus system, Duke University, Wake Forest University, and others. The network provides high-speed Internet service, access to Internet2 and the national research and

The Role of Standards in the Grid

Continued from page 13.

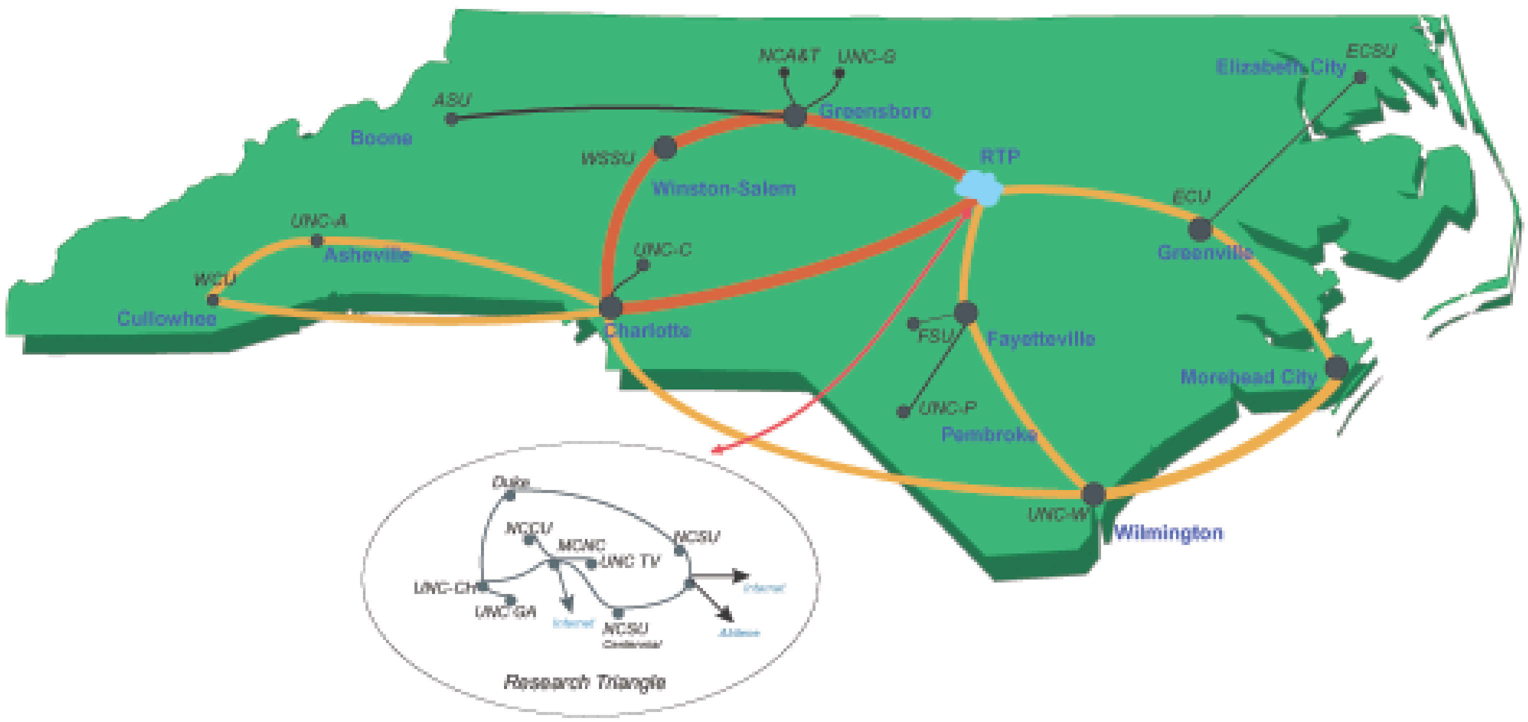
directs the Argonne Distributed Systems Laboratory, which designs and develops the Globus Toolkit and standards on which it is based. The toolkit has been adopted as the central component in Grid solutions offered by IBM, HP, Oracle and other major IT companies.

Author Contact Information

Lee Liming
Argonne National Laboratory
E-mail: liming@mcs.anl.gov

Ian T. Foster
Argonne National Laboratory
E-mail: foster@mcs.anl.gov

North Carolina Research & Education Network (NCREN)



education Abilene network, and interactive, near broadcast-quality video conferencing for distance-learning classes.

The N.C. Bioinformatics Grid

Scientific communities, with exponentially increasing storage and resource needs, are driving the development of grid computing frameworks to support the next generation of innovation. Biology and life sciences researchers, historically not heavy users of high-performance computing, are now at the leading edge of this evolution as the sequencing of entire genomes has unlocked a new horizon of opportunity. The availability of massive compilations of genomic and related data is merging biology with information science. Storage and management of these data sets will require systems capable of managing petabytes of data, and analysis and modeling will require high-performance computing capabilities.

With enhanced capabilities to address computational and data requirements, grid computing is an ideal solution to the needs of life sciences researchers. Also, the transparent and seamless access to compute and data storage resources, as if they are located on a single computer, makes grid solutions an even more compelling fit.

MCNC and the North Carolina Biotechnology Center's Genomics and Bioinformatics Consortium, in collaboration with IBM, launched the N.C. BioGrid in 2001 as one of the nation's first grid test beds for computing, data storage and networking resources for life sciences research. The N.C. Bioinformatics Consortium includes more than 80 organizations representing academia, business and industry. Members include the University of North Carolina's 16-campus system, Duke University, Wake Forest University, GlaxoSmithKline Inc., IBM, the Research Triangle Institute, SAS Institute, Biogen, the National Institute of Environmental Health Sciences, and the U.S Environmental Protection Agency.

> "As the grid virtual team has learned more about grid computing and delved further into the NC BioGrid and NCREN, we have come to view the BioGrid as one of NCREN's most innovative applications and the one area that can demonstrate networking's greatest potential."
>
> *Wayne Clark*
> *Networking Services Architect*

Work is being conducted to test software for a better understanding of the issues associated with storage, analysis and movement of large bioinformatics data sets in a high-speed networked environment. The objective is to enable participants to share data and computing resources, thus eliminating the need for costly duplication of data sets and computing resources at each institution.

Currently, the test bed involves resources from four organizations – the University of North Carolina at Chapel Hill, North Carolina State University, Duke University and MCNC.

continues on page 16

North Carolina Statewide Grid

Continued from page 15.

In planning the BioGrid test bed, a number of key objectives were identified:

- The grid must span multiple administrative domains and allow for the independent management of resources.
- A diverse set of hardware and operating system platforms should be supported in the grid.
- Systems that participate in the grid must be configured to meet a superset of the security policies of each organization.
- Data files should be organized in a global namespace so that they can be accessed at the file system level with consistent pathnames throughout the grid.
- The grid must minimize network traffic through data caching and replication.
- For sensitive data being transported across the network, the grid must provide facilities for data encryption and integrity checks.
- The grid requires a "meta-scheduler" that can intelligently and transparently select compute resources capable of running a user's job.
- The grid creates a uniform name space so that resources can be addressed consistently across the grid.

BioGrid Middleware and Supporting Technologies

Perhaps the biggest challenge for the N.C. BioGrid was to identify the appropriate grid middleware. This led to an evaluation of numerous grid platforms and testing a mix of solutions. A hybrid solution of multiple grid middleware platforms was developed, working with technologies that have a relationship with or roadmap to Open Source Grid Architecture (OSGA) standards. Currently, the following grid platforms are deployed:

Globus Toolkit 2.4 – for core grid functionality such as job scheduling across administrative domains, a resource registry, and a framework for developing grid-aware applications.

Avaki Data Grid 3.0 — to provide a globally available file system using a global namespace.

In addition, we are working with the following supporting technologies:

Platform LSF — for job scheduling on clusters and large SMP servers.

Sun Grid Engine — for job scheduling on clusters.

Sun ONE Directory Server — LDAP infrastructure for managing user accounts.

CHEF – a framework for building grid-aware collaborative portals.

MyProxy – an online repository that enables remote management of grid credentials.

BioGrid Applications

The first prototype application selected for the N.C. BioGrid was NCBI BLAST. This tool is widely used in the bioinformatics research community to search for similarities between candidate proteomic or nucleotide sequences and target genomes.

IBM worked with MCNC and its university partners to develop more sophisticated grid applications. In an Extreme Blue project conducted during 2003, IBM teamed a group of four student interns with mentors to build a grid-enabled interface to the BioPerl libraries to address the computational needs of the Fungal Genomics Lab at N.C. State University. **Results that took one to two weeks using a single system are now produced in near real time on the BioGrid**. The Fungal Genomics Lab has also integrated one of its clusters with the BioGrid test bed.

In a second example, IBM and MCNC worked with researchers at the University of North Carolina at Chapel Hill to build a grid-enabled drug discovery application that screens candidate chemical compounds for biological activity. This is accomplished by performing a parameter space study to produce a training set to develop a model. The model is then applied to other data. **Work that previously took a month is now accomplished in a single day.**

The MCNC Enterprise Grid

To address multiple disciplines beyond the original scope of the N.C. BioGrid test bed, MCNC has reconfigured its compute, storage, data and application resources into a grid architecture – the MCNC Enterprise Grid.

Biology research is an application for grid, but the test bed infrastructure is now evolving to address multiple research disciplines and applications. As campus infrastructures are moving to grid frameworks, the development of the MCNC Enterprise Grid is a step towards the "grid of grids" concept. As computing and storage clusters evolve into grids, they will be interconnected into larger grids that will cross multiple organizational boundaries (firewalls), such as through the North Carolina statewide grid.

The initial launch of the MCNC Enterprise Grid in October 2003 included two high-performance computing systems:

- A 64-node Massively Parallel Processor (MPP), distributed memory IBM Cluster with a total of 128 Intel 2.8-GHz, 32-bit CPU's running RedHat Linux

- A Symmetric Multi-Processor (SMP), shared memory SGI server with 32 Intel 1.3-GHz, 64-bit CPU's running a variant of the Advanced Server edition of RedHat Linux

A combination of direct-attached and network-attached (Network Appliance) disks complement the computer systems with over 10 Terabytes of storage. Gigabit Ethernet (Cisco), Infiniband (Topspin), and Fiber Channel (IBM) comprise the varied technologies used for interconnection and switching between the compute and storage nodes.

In addition to academic use, North Carolina commercial organizations may also use the MCNC Enterprise Grid as a fee-based service for research purposes. Charges are based on a per CPU-Hour basis or negotiated rates for dedicated access. Services include up to 2 gigabytes of home directory space and a selection of software packages.

Researchers are using the grid resources for a variety of tasks, including scientific modeling and analysis. The on-demand utility computing services model allows customers to pay only for what they need, when they need it. The shared resources reduce the requirement for large investments in high performance computing hardware and support staff at businesses and universities.

MCNC's Enterprise Grid supplements the N.C. BioGrid test bed. It is a resource for the development of a new Grid Technology Evaluation Center and the North Carolina statewide grid.

The Grid Technology Evaluation Center

The Grid Technology Evaluation Center (GTEC) is another development that emerged from MCNC's experience gained with the N.C. BioGrid. MCNC is working with commercial industry partners to develop the center, which will further address the challenges associated with moving grid technologies from the research lab and test bed environment to core enterprise infrastructure and of the emerging "Next Generation Internet" that delivers a new generation of digital consumer services.

The on-demand utility computing services model allows customers to pay only for what they need, when they need it.

The GTEC will facilitate, enhance, enable, and expedite the development and deployment of grid computing infrastructure and services through:

- A test bed supporting integration, experimentation, development, and training in grid deployment across multiple research disciplines and applications.
- A platform for the integration, testing, and development of grid infrastructure, middleware, applications, API's and other grid-related technologies.
- A facility to support grid development activities.

GTEC services will include application benchmarking, interoperability verification, systems integration (including integration with legacy systems), and operational training.

MCNC's Related Grid Research & Development

As an emerging technology, it will take years before the ubiquitous use of grid computing on MCNC's statewide network is realized. As early grid technology adopters and active participants in standards bodies, MCNC has been able to identify the challenges in deploying, operating, and scaling a grid infrastructure beyond the test bed phase. MCNC Research & Development Institute's grid-related research focuses on filling the gaps in existing grid-ware to address these challenges, as shown in the accompanying illustration.

Some of the research efforts include:

Grid Middleware

GridIR: A grid-based information retrieval system that provides a scalable framework to uniformly search and retrieve public and private diverse data across the grid while allowing local control on the data. MCNC is also actively involved in the Global Grid Forum (GGF), forming and participating in the GGF GridIR working group to standardize the interfaces for providing information retrieval capability in a grid environment.

GridScope: An effort to build a grid monitoring and tracking tool that presents a logical view of the interactions within and between grid applications and captures the grid interactions as well.

Cluster-on-Demand (COD): A system to enable rapid, automated, on-the-fly partitioning of a physical cluster into multiple independent virtual clusters.

Security Infrastructure

Most organizations today have firewalls around their organizational computer resources to protect their sensitive and proprietary data. Grid topologies span multiple administrative domains with autonomous security mechanisms. Unlike the Internet, the grid allows an outsider complete access to the resource, thus increasing the risk associated with it. The central idea of grid computing to enable sharing of resources across existing organizational and geographical boundaries makes it difficult to use existing security mechanisms such as firewalls on the grid.

Though organizations may be willing to share resources and data with others for collaborative or monetary

continues on page 18

North Carolina Statewide Grid

Continued from page 17.

reasons, information assurance must be guaranteed for participation in a grid environment. It is imperative to bridge the gap between different security mechanisms, while providing local autonomy.

Joint Control of Virtual Organizations (JoVO): JoVO seeks to address difficulty of mapping into virtual organizations some of the typical social and political arrangements that are associated with shared resources that need joint control. The objective of JoVO is to develop a scalable, reliable, distributed identity, authentication and authorization infrastructure to facilitate secure collaborations in grid environments. The perceived JoVO framework will enable multiple parties to form a virtual coalition with jointly agreed and enforceable rules to enable timely information sharing and collaborative processing. The joint control of identity, attributes and access control policy is achieved through the use of threshold-based certification authorities. The framework is a public key infrastructure (PKI) that is both fault and intrusion tolerant.

The technical approach to JoVO is based on MCNC's completed research project funded by the Defense Advanced Research Projects Agency (DARPA).

SITAR: The need to provide information assurances for data and applications necessitates the need for intrusion and fault tolerant security capabilities. SITAR (Scalable Intrusion-Tolerant Architecture) is designed to ensure that critical services and applications remain operational, even while under attack.

SITAR, a completed DARPA-funded project, is an extensible framework that incorporates the fault tolerant concepts of redundancy, diversity, and ballot voting along with adaptive reconfiguration and proactive monitoring for extending fault tolerance to distributed services. Its fault tolerance approach focuses on detecting and mitigating the effects of known and unknown intruder attacks that attempt to interrupt service availability.

Grid represents a fundamental turning of the technology crank and is the next big thing in the evolution of the Internet.

Network Provisioning

High-speed, on-demand, application-initiated provisioning of bandwidth will improve the efficiency and reduce the latency in a grid network. In January, MCNC announced the successful demonstration of an optical network provisioning protocol to enable more efficient computing applications. The demonstration of the Just-in-Time (JIT) protocol for provisioning and managing light path connections in the all-optical Advanced Technology Demonstration Network (ATDnet) in Washington, D.C., confirmed the viability of user-initiated, ultra-fast provisioning of all-optical network connections. The light paths linked host systems at the U.S. Department of Defense's Laboratory for Telecommunications Sciences, the Naval Research Laboratory's Center for Computational Science and the Defense Intelligence Agency.

With JIT, optical connections can be provisioned between sites in a few milliseconds through microelectromechanical switches, and in a few microseconds when faster photonic switches are deployed.

JIT research was partially funded by NASA and supported by the Advanced Research and Development Activity, a Department of Defense research and development community.

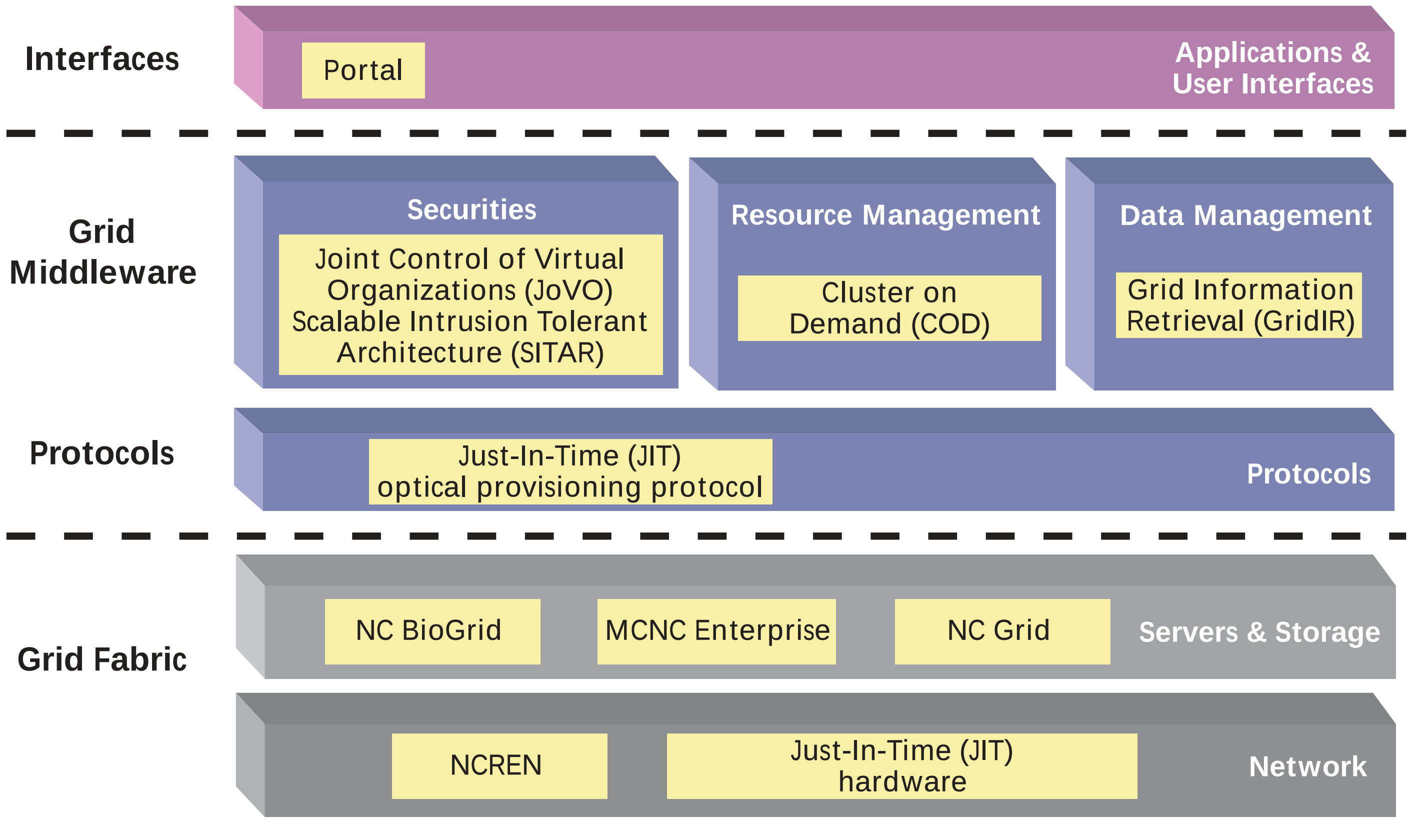
Reaching the Potential of Grid Computing

In developing the N.C. BioGrid, MCNC created the foundation for a statewide grid infrastructure that will support on-demand access to resources and services. The MCNC enterprise grid, which provides core computing and storage resources to researchers, serves as a model for integration with and access to inter-domain and global grids. The GTEC provides a platform for the development and integration of grid infrastructure, applications and services.

Grid represents a fundamental turning of the technology crank and is the next big thing in the evolution of the Internet. MCNC is moving aggressively to accelerate the integration of grid technology into both the fabric of North Carolina's Internet infrastructure and into the fabric of the new economy.

About MCNC

MCNC is a private, independent, non-profit corporation established in 1980 to advance technology-led economic development and job creation throughout North Carolina. MCNC Research & Development Institute develops new technologies through its own initiatives and as a research partner for private industry and the U.S. government, conducting advanced and applied research across a broad technology spectrum, including microsystems, flexible electronics, sensor development, signal electronics,



MCNC Grid Project Stack

wireless systems, microfabrication, high-speed secure networks and grid computing. MCNC Grid Computing & Networking Services delivers advanced communications resources statewide to more than 180 public and private institutions. MCNC Ventures provides early-stage funding and assistance to entrepreneurial start-up companies. The MCNC family of companies is located in North Carolina's Research Triangle Park. For more information, please visit www.mcnc.org.

About the Authors

Phil Emer is a senior member of MCNC Grid Computing & Networking Services' Advance Technologies Group technical staff and program director for The Grid Technology Evaluation Center. He has spent nearly 15 years working at the intersections of networking, research and academia. He was the chief architect of the N.C. BioGrid, managing the development of a heterogeneous, multi-institutional, grid test bed that spans MCNC, N.C. State University, Duke University, and the University of North Carolina at Chapel Hill.

Chuck Kesler is a program manager for MCNC's Grid Computing & Networking Services' grid deployment and data center services. He provides technical architecture and project management for MCNC's grid computing and hosting initiatives. His activities have included spearheading the deployment of the North Carolina BioGrid test bed and leading a collaborative grid infrastructure working group that includes representatives from the local university community.

Lavanya Ramakrishnan is a research engineer for MCNC Research & Development Institute. She is currently involved with various grid and security projects, including the development of a security infrastructure for grid applications. She is also involved with designing portal-based interfaces to grid functionality and development, testing and evaluating the use of various grid middleware to be deployed on several grid test beds, including the NASA-funded Virtual Collaborative Center and N.C. BioGrid. She serves as a senior leader for the 'Cluster On Demand' project, a collaboration with Duke University funded by the NSF Middleware Initiative, to develop a grid service for dynamic virtual clusters. She is also actively involved in the security working groups at the Global Grid Forum (GGF).

Contact Information

Scott Yates
Director of Corporate Communications
MCNC
P.O. Box 13910
3021 Cornwallis Road
Research Triangle Park, NC 27709

PH : (919)248-1907
FAX : (866)773-5617
E-mail : syates@mcnc.org

TeraGrid Software Strategy: E Pluribus Unum

By Patricia Kovatch, San Diego Supercomputer Center

TeraGrid: Background and Development

"Build and deploy the world's fastest distributed infrastructure for open scientific research:" this is the mission of the TeraGrid project. It is a ~$100M multi-year project, sponsored by the National Science Foundation (NSF). It's the first joint multi-site supercomputing effort featuring over 24 TeraFlops (TF) of initial combined compute power (13 TF of Itanium2), one PetaByte of online storage and a 40 Gigabit per second backbone. Additional sites are in the process of adding their resources to the TeraGrid. A grid infrastructure was developed and deployed to help scientists and researchers make use of these resources. Cosmology, weather and geophysics applications make up the initial allocations on the TeraGrid. Coordinating the efforts to manage the sheer amount of software, operations and user requests has been an interesting challenge.

The NSF vision, as given in the Request For Proposal (RFP) for the TeraGrid project, reads, "NSF seeks to open a pathway to future computing, communications and information environments by creating a very large-scale system that is part of the rapidly expanding computational grid." The goals of the TeraGrid are to "enable new science by offering new capabilities, build an extensible grid that can be grown and copied, and provide an evolutionary pathway for current users." The solicitation specified that the proposals must include distributed, multi-site facilities with single site and "grid enabled" high end compute capabilities connected via ultra high-speed networks. A distributed storage system with both online and archival storage capabilities was required. Remote visualization and a production-level quality of service were necessary components.

Many Major Research Equipment (MRE) projects have common needs: geographically distributed instruments, Terabytes to PetaBytes of data, data unification and sharing between multiple formats and remote sites, high end computation (simulations, analysis and data mining), presentation of results (visualization, virtual reality, etc.). Examples of these projects include:

Coordinating the efforts to manage the sheer amount of software, operations and user requests has been an interesting challenge.

- Atacama Large Millimeter Array (ALMA)
- EarthScope (Structure and Evolution of the North American Continent)
- IceCube (High Energy Neutrino Detector)
- Laser Interferometer Gravitational Wave Observatory (LIGO)
- Network for Earthquake Engineering Simulation (NEES)

Considering the needs of these projects, we sought to create a seamless environment for the geographically distributed scientists of the TeraGrid. We decided to minimize differences in hardware and software. We researched and tested ways to make resources available equally from each site. We defined a unified user support operation and we developed an automatic reporting and testing infrastructure to guarantee that a certain service level of resources would be available to our scientists and researchers.

January 1, 2004 marked the first day of production for the four founding sites of the TeraGrid. To make this happen, these four diverse sites had to define and implement common goals. Thirteen working groups were ordained to flesh out the vision and make detailed implementation plans. The working groups consisted of Clusters, Grid, Networks, Performance Evaluation, User Services, Data, Operations, Applications, Visualization, External Relations, Interoperability, Account Management, and Security.

TeraGrid: Software Management Strategies

The basic TeraGrid software stack consists of:

- A base operating system
- Applicable drivers (for example, the SDSC Itanium2 cluster has *Myrinet*, *Qlogic* (fibre channel), and *SysKonnect Gigabit Ethernet* drivers)
- Compilers (*GNU* and *Intel* for TeraGrid Itanium2 clusters)
- Message passing interfaces (*MPICH, MPICH-GM* (for Myrinet on Itanium2)
- Numerical libraries (*FFTw, ScaLAPACK*, etc.)

TeraGrid Founding Sites
San Diego Supercomputer Center (SDSC)
National Computational Science Alliance (NCSA)
Argonne National Laboratory (ANL)
California Institute of Technology (Caltech)

New Sites
Pittsburgh Supercomputer Center (PSC)
Indiana University
Purdue University
Texas Advanced Computing Center (TACC)
Oak Ridge National Laboratories (ORNL)

Corporate Partners
IBM, Intel, Qwest, Sun, Myricom, Oracle

- Data formatting tools (*HDF4* and *HDF5*)
- Archival storage and data collection access
- Database clients
- Parallel filesystem daemons
- Resource manager (*Portable Batch System (PBS)* on Itanium2)
- Globus Grid infrastructure

Many other applications, such as the *Storage Resource Broker (SRB)*, help the scientists manage their data sets for their research. In total, there are over 100 separate pieces of software. Keeping all of these pieces of software up-to-date and interoperable at many different sites can be unwieldy, but there are several techniques TeraGrid employed to manage all of these pieces.

Use of a Global Software Repository Each site has representatives responsible for parts of the software stack who are tasked with watching for updates to the software (including security), testing these updates and checking in the software, accompanying configuration files and build instructions to the global software repository. Other sites then check out the programs and files from the global software repository and install them locally. TeraGrid currently uses an open source repository tool called *Concurrent Versions Systems (CVS)* to manage the checking in and checking out of software versions of all of our software, from the kernel to applications. Version changes are typically applied during one of the regularly coordinated, scheduled (weekly) Preventive Maintenance (PM) periods which are staggered between the sites so not all resources are unavailable at the same time. This scheduling strategy allows software (and hardware) upgrades to happen in a timely manner. Updates to the software stack due to security concerns such as root exploits happen immediately. Policies and procedures are in place to prevent, protect, detect and handle security incidents.

Maintenance of Nodes The nodes are maintained using a standard suite of tools for automatic and repeatable installation and updating software, automatically parsing the logs, etc. Being able to update and check the software on all the nodes (as well as install the nodes) is critical to maintaining systems with multiple nodes.

Consistent User Environment To abstract away the local dependencies at each site, agreed upon environmental variables are used to make users' scripts portable. This enables users to experience a consistent environment regardless of which site is accessed, an attribute called TeraGrid "roaming". There are more than 20 variables defined for scratch and parallel filesystems, library paths, *MPICH*, etc. The software that manages setting the user environment variables is called *Softenv*. More information about *Softenv* is available at http://www-unix.mcs.anl.gov/systems/software/msys/ .

Sophisticated Monitoring Tools One "at-a-glance" view of each site's software versions and basic software functionalities helps everyone view the status of the system immediately. This "*Inca Test Harness and Reporting Framework*" is a flexible software infrastructure for automated testing, verification and monitoring of the TeraGrid software stack and environment. *Inca* is composed of a set of "reporters" that interact with the system and report status information, a "harness" that provides basic control of the reporters and collection of information, and "clients" that provide a web interface to the information collected by the reporters. Currently over 900 pieces of information about the software and environment are checked.

A reporter is a self-contained pluggable component implemented as a script or executable that performs a test, benchmark or query and outputs a result in XML. For example, a reporter can output package version information or test whether a grid service is up and available. Reporters can be written in any language and Perl and Python API helper libraries are available to ease the process of writing reporters. Currently, there are more than 100 reporters written with these API's. The execution details of the reporters (frequency, inputs, etc.) are handled by the harness. The harness contains a set of daemons which manage the distributed execution of the reporters, collect the

continues on page 22

TeraGrid Software Strategy

Continued from page 21.

data to a central location and publish the data into an information service such as *MDS2*. Clients compare the resource information collected by *Inca* against a representation of the software stack and environment and display it on a TeraGrid web status page. A version of the *Inca Test Harness and Reporting Framework* is available for download at http://tech.teragrid.org/inca.

Software Change Request Process Anyone can submit a request to add or update software to the TeraGrid software stack by submitting a software change request. An organized plan for testing, deployment and integration of the software is developed and implemented. Test nodes at each site are connected and help with the interoperability testing.

Use of the Globus Grid Toolkit The toolkit helps create a geographically distributed infrastructure between the sites. *Gx-map* and *CACL* make it easier to manage this infrastructure. *CACL* is an Open SSL based Certificate Authority (CA) CLient system that issues digital certificates. It automatically authenticates the identity of the user using the username and password. The encrypted request is submitted to the CA daemon, which decrypts it and authenticates the user in the same way that ordinary login authentication is done. It then either issues a certificate or a rejection notice to the CA CLient program. If the certificate is issued, it is automatically placed in the user's home directory. There is a program that the CA administrator can use to revoke certificates and create server or service certificates. These commands can only be run by a system administrator on the machine that provides the CA services. More information is available at http://www.sdsc.edu/CA.

Gx-map allows users to add their certificate automatically to the grid-mapfile. This file maps a specified Distinguished Name (DN) to a Unix account name. *Gx-map* propagates these changes automatically between systems. An auxiliary tool called *gx-check-cacl-index* automatically checks for new user certificates issued or revoked by *CACL* and invokes the *gx-map* command to request the appropriate updates. With these systems, a user can obtain a digital certificate and update the grid-mapfile entries on a number of systems without manual intervention from systems administrators. This software can be found at http://www.sdsc.edu/~kst/gx-map.

To make it easy for scientists and researchers to get started on the TeraGrid, there is one location that gives information about getting accounts and allocations: http://accounts.teragrid.org

TeraGrid: Account Management and Accounting

To make it easy for scientists and researchers to get started on the TeraGrid, there is one location that gives information about getting accounts and allocations: http://accounts.teragrid.org. Account requests funnel to the national, peer-reviewed allocation committees. The National Resource Allocation Committee (NRAC), Partnership Resource Allocation Committee (PRAC) and Alliance Allocations Board (AAB) meet several times a year and allocate over ten million Service Units per year. A Service Unit (SU) is approximately one Central Processing Unit (CPU) hour. Most allocations last for one year, though multi-year allocations are now available to scientists with multi-year projects. Start-up allocations can be granted by the Development Allocations Committee (DAC) for up to 30,000 SUs. Once allocations are approved, the project and account notifications are transported between the remote sites and the TeraGrid Central Database (Postgresql) located at SDSC via the *Account Management Information Exchange (AMIE)* system. *AMIE* uses an XML format to facilitate import and export between each site's local database and central database. Soon Perl scripts will automatically transport and accept account addition and deletion requests and allocation usage statistics from SDSC to the other sites.

When a user submits a job to a TeraGrid machine, a wrapper script around the job submission commands (*PBS* or *Globus*) checks to see whether the account has an allocation. The job runs if the account has enough allocation to fulfill the job request. After the job completes, the job information is logged. Then the *TGAccounting package*, developed at SDSC, converts the resource manager (*PBS*) and *Globus* logs into an XML format compatible with the Global Grid Forum (GGF) Usage Record Format and National Middleware Initiative accounting protocols. These scripts run nightly. The XML is then parsed and imported into records in the local site's database and finally transported via *AMIE* to the TeraGrid Central Database where the allocation is deducted. Soon the allocations will be fungible between sites.

TeraGrid: Operations and Unified Help Desk

Though the TeraGrid resources span several sites, the TeraGrid Help Desk appears as one entity to the scientists and researchers who use it. There is one phone number, one e-mail address (help@teragrid.org) and one website (http://www.teragrid.org) presented to the users. To support this, there is one ticket system that accepts the trouble tickets via e-mail or web input. This ticket system has been specially tailored by NCSA to quantify and track tickets for the TeraGrid. It automatically notifies the representatives from each site when new tickets arrive. The tickets can be assigned to a specific site or all sites depending on the request in the ticket. Site representatives update tickets via a webpage interface to the ticket system database.

Both NCSA and SDSC have round-the-clock operations centers and trade off the monitoring to provide coverage on a 24x7 basis. These centers have procedures for contacting people during off-hours for support as needed.

Each TeraGrid site has an operations center with monitors dedicated to displaying the status of the machines. *Clumon*, software developed at NCSA, monitors performance metrics from the nodes and queues and correlates data and other queue and job information from the resource manager and scheduler. It provides a near-real time, graphical display about the health of the system and the processes running on it. Visit http://clumon.ncsa.uiuc.edu for more information. Another monitoring tool called *Ganglia* is also used.

TeraGrid: Collaboration

Each site brings different expertise to the TeraGrid. Each site has provided critical insight, ideas and implementation effort to the TeraGrid project. Each site has contributed essential software development, testing and troubleshooting. The sites have worked closely together to develop the design and implementation of the TeraGrid facilitated by weekly conference calls and periodic face-to-face meetings.

TeraGrid: Futures

Research and demonstration of new technologies that improve upon the current implementation are continually being conducted. Specific areas include automatic scheduling of resources at multiple sites (metascheduling), sharing home directories over the wide area (parallel filesystems mounted over WAN) and remote manipulation of storage facilities (Fibre Channel over IP protocol). In all of these cases, prototype programs have been demonstrated at the Supercomputing conferences. For more information on these demonstrations, please visit http://www.sdsc.edu/~pkovatch/new-tech .

As new sites join on, new resources are available and new interoperability and scalability issues are identified. Growth will continue to provide new challenges and rewards to the scientists and researchers making use of the ever-expanding TeraGrid.

About the Author

Patricia Kovatch is the Manager of the High Performance Computing Production Archival Storage Systems at the San Diego Supercomputer Center (SDSC). She is SDSC's lead for the Cluster and Grid Working Group. She also serves on the Data and Performance Evaluation Working Groups. She has worked on 512 processor commodity clusters with Myrinet and currently Itanium2 clusters for the TeraGrid project.

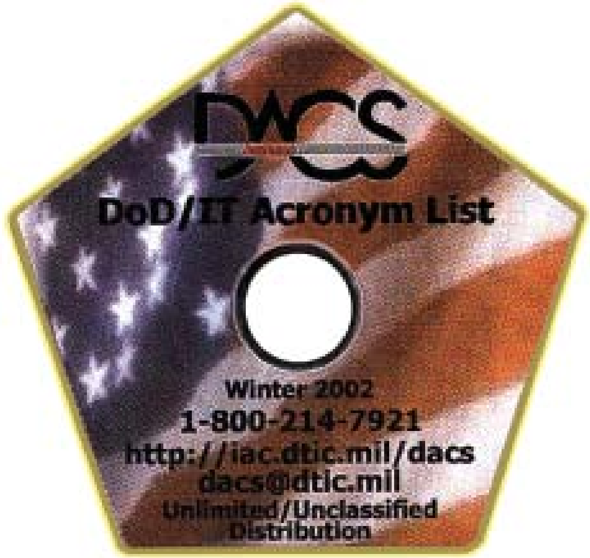
Author Contact Information

Patricia Kovatch
Manager, HPC Systems and Archival Storage
San Diego Supercomputer Center
PH: (858)822-5441
E-mail: pkovatch@sdsc.edu


Great Software Tools
from the DACS!
Return On Investment from
Software Process Improvement
Products CD
New XML Database
Included!
ROI
The DACS
Software Reliabili
Sourcebook
ROI/SPI CD:
• Technical Report &
Spreadsheet Model
• ROI Database
• Cost Benefit Analysis
SW Reliability
Sourcebook CD:
• Covers the Software
Life Cycle
• Hundreds of Hotlinks &
Embedded Documents
http://iac.dtic.mil
Purchase these tools at the DACS Store
www.dacsstore.com
or contact
800-214-7912 dacs@dtic.mil

The first 50 people to send in a completed survey will receive a FREE DoD/IT Acronym CD from the DACS.

This valuable CD-ROM contains over 9,000 Department of Defense and Information Technology acronyms. There are hundreds of acronym lists available but none are as well done as this CD AND specifically targeted towards DoD and Information Technology. This unique-shaped CD-ROM plays in your computer's regular, hub-mounted, CD drive. You'll use this great resource over and over again. It's FREE, just for filling out our brief survey on the next page!



▼ Fold Here ▼

Data & Analysis Center for Software (DACS)

http://iac.dtic.mil/dacs/

▼ Fold Here ▼

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 120 ROME NY

POSTAGE WILL BE PAID BY ADDRESSEE

DATA & ANALYSIS CENTER FOR SOFTWARE
775 DAEDALIAN DR
ROME NY 13449-0139

Software Tech News Subscriber Survey

1. Which volume of the Software Tech News did you receive? STN 7:1 Grid Computing

2. When did you receive the newsletter? *(month/year)* ____________

3. How satisfied were you with the CONTENT of the newsletter? (Article Quality)

❏ Very Satisfied ❏ Satisfied ❏ Neither Satisfied nor Dissatisfied ❏ Dissatisfied ❏ Very Dissatisfied

4. How satisfied were you with the APPEARANCE of the newsletter?

❏ Very Satisfied ❏ Satisfied ❏ Neither Satisfied nor Dissatisfied ❏ Dissatisfied ❏ Very Dissatisfied

5. How satisfied were you with the OVERALL QUALITY of the newsletter?

❏ Very Satisfied ❏ Satisfied ❏ Neither Satisfied nor Dissatisfied ❏ Dissatisfied ❏ Very Dissatisfied

6. How satisfied were you with the ACCURACY of the address on the newsletter?

❏ Very Satisfied ❏ Satisfied ❏ Neither Satisfied nor Dissatisfied ❏ Dissatisfied ❏ Very Dissatisfied

7. Approximately how much of the newsletter do you read?

❏ The entire issue ❏ Most of the content ❏ About half the content ❏ Briefly Skimmed ❏ Didn't Read

8. Would you read this newsletter in an E-mail newsletter format?

❏ Definitely ❏ Probably ❏ Not Sure ❏ Probably Not ❏ Definitely Not

9. How did you request the product or service?

❏ Phone Call ❏ E-mail ❏ DACS Website ❏ Subscription Form Other ____________

10. Would you recommend the DoD Software Tech News to a colleague?

❏ Definitely ❏ Probably ❏ Not Sure ❏ Probably Not ❏ Definitely Not

11. What topics would you like to see this newsletter devoted to?

____

Comments (Optional)

____

____

Contact Information (Optional)*

Name: Position/Title:

Organization: Office Symbol:

Address:

City: State: Zip Code:

Country: E-mail:

Telephone: Fax:

Functional Role:

Organization Type: ❏ Air Force ❏ Army ❏ Navy ❏ Other DoD ____________

❏ Commercial ❏ Non-Profit ❏ Non-US ❏ US Government ❏ FFR&D ❏ Other ____________

*Note: You must give us your address to receive the CD.

About the DoD Software Tech News

Article Reproduction

Images and information presented in these articles may be reproduced as long as the following message is noted:

"This article was originally printed in the *DoD Software Tech News*, Vol. 7, No. 1. Requests for copies of the referenced newsletter may be submitted to the following address:

Philip King, Editor
Data & Analysis Center for Software
P.O. Box 1400
Rome, NY 13442-1400

Phone: 800-214-7921
Fax: 315-334-4964
E-mail: news-editor@dacs.dtic.mil

An archive of past newsletters is available at www.SoftwareTechNews.com.

In addition to this print message, we ask that you send us three copies of any document that references any article appearing in the *DoD Software Tech News*.

About This Publication:

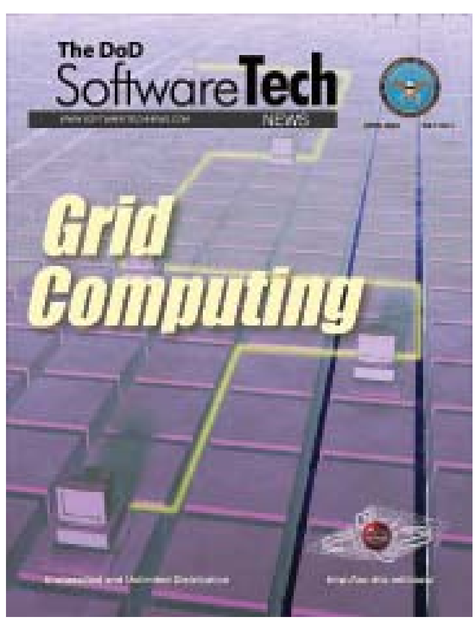
The *DoD Software Tech News* is published quarterly by the Data & Analysis Center for Software (DACS). The DACS is a DoD sponsored Information Analysis Center (IAC), administratively managed by the Defense Technical Information Center (DTIC) under the Defense Information Systems Agency (DISA). The DACS is technically managed by Air Force Research Laboratory, Rome, NY and operated by ITT Industries, Advanced Engineering and Sciences Division.

Data & Analysis Center for Software
http://iac.dtic.mil/dacs/

DACS
P.O. Box 1400
Rome, NY 13442-1400

Phone: 800-214-7921
Fax: 315-334-4964
E-mail: dacs@dtic.mil
URL: http://iac.dtic.mil/dacs/

Cover Design by Joe Barbaccia, ITT Industries



To Subscribe to this Publication Contact:

Phone: 800-214-7921
Fax: 315-334-4964

E-mail: news-editor@dacs.dtic.mil
Web: www.dacs.dtic.mil/forms/regform.shtml

Distribution Statement:

Unclassified and Unlimited

STN Vol. 7, No. 1

In This Issue

Philip King, STN Editor
Data & Analysis Center for Software
P.O. Box 1400
Rome, NY 13442-1400

Phone: (800) 214-7921
Fax: (315) 334-4964
E-mail: news-editor@dacs.dtic.mil

Data & Analysis Center for Software
P.O. Box 1400
Rome, NY 13442-1400

Return Service Requested

PRSRT STD
U.S. Postage
P A I D
Permit #566
UTICA, NY